F<sub>1</sub> macro scores (100 test points, GPT-3 fine tuned on SMILES)  $\eta$  multiplier=0.02  $\eta$  multiplier=0.05  $\eta$  multiplier=0.1  $\eta$  multiplier=0.2 epochs=1 0.0 epochs=2 epochs=4 0.5 0.0 epochs=8 0.5 ada babbage random 0.0 100 200 100 200 100 200 100 200 number training points number training points number training points number training points